

Algorithmisch Rekursive Sequenzanalyse 2.0 – Beispiel mit kritischer Reliabilitätsprüfung

Paul Koop

November 2024

post@paul-koop.org

Zusammenfassung:

Mit der Methode der Algorithmisch Rekursiven Sequenzanalyse soll das transkribierte Protokoll eines Verkaufsgesprächs analysiert werden. Es wird unterstellt, dass man, um die Intentionen der Sprecher zu rekonstruieren, zunächst klären muss, was die Interaktionen im Sinne von Regeln einer probabilistischen kontextfreien Grammatik bedeuten könnten. Nach einer Voruntersuchung zur Hypothesenbildung zu Verkaufsgesprächen sollen die Interaktionen orientiert an, aber unabhängig von der hypothetisch gebildeten Grammatik so untersucht werden, dass aus den Interaktionen eine terminale Zeichenkette abgeleitet wird, wobei jedem Interakt ein Symbol für ein Terminalzeichen zugeordnet wird. Anschließend soll aus der Terminalzeichenkette eine Grammatik induziert werden. Diese empirische Terminalzeichenkette wird auf Wohlgeformtheit in Bezug auf die induzierte Grammatik geprüft. Danach werden aus dieser Grammatik künstliche Terminalzeichenketten generiert, die auf signifikante Übereinstimmungen mit der empirisch ermittelten Grammatik untersucht werden. In einem weiteren Schritt kann die Grammatik einem (hybriden) Multiagentensystem übergeben werden, das sich bei der Simulation von Verkaufsgesprächen an den Regeln und dem Protokoll der Grammatik orientiert.

Reliabilität der Kodierung: Kritische Prüfung und Bereinigung

Die Grundlage jeder weiteren Analyse ist die Zuordnung der 59 Interakte zu 12 Kategorien durch zwei unabhängige Kodierer (Kodierer 1 und Kodierer 2). Die ursprünglich verwendete Formel

$$R_{\text{ars}} = \frac{N \cdot Z}{I} = \frac{2 \cdot 35}{118} = 0,59$$

misst die beobachtete Übereinstimmung ($p_o = 0,59$), korrigiert diese jedoch **nicht** um den Anteil zufällig erwarteter Übereinstimmungen. Bei 12 gleichwahrscheinlichen Kategorien liegt die Zufallserwartung bei

$$p_e = \frac{1}{12} \approx 0,0833.$$

Der korrigierte Koeffizient (Cohens Kappa) berechnet sich daher als

$$\kappa = \frac{p_o - p_e}{1 - p_e} = \frac{0,59 - 0,0833}{1 - 0,0833} \approx 0,55.$$

Mit $\kappa \approx 0,55$ liegt die bereinigte Reliabilität nur noch im unteren Bereich von mittel (nach Landis & Koch) und **unterschreitet** die üblichen Akzeptanzgrenzen von $\kappa \geq 0,70$ (explorativ) bzw. $\kappa \geq 0,80$ (konfirmatorisch). Die ursprüngliche Formel überschätzt die wahre Übereinstimmung systematisch. Für eine belastbare Inhaltsanalyse ist eine Reliabilität von 0,55 nicht ausreichend.

Kreuztabellen und ergänzende Chi-Quadrat-Tests

Die Chi-Quadrat-Werte beziehen sich auf **Anpassungstests für jeden Kodierer einzeln**, nicht auf die Übereinstimmung zwischen beiden. Sie prüfen, ob die beobachtete Verteilung der Kategorien von einer erwarteten Verteilung abweicht.

Table 1: Chi-Quadrat-Tests pro Kodierer

Kodierer	χ^2	df	Asymp. Sig.	vorab gesetztes α
1	2,6	6	0,86	0,9
2	2,0	5	0,85	0,9

Die hohen asymptotischen Signifikanzen ($p > 0,05$) zeigen, dass für keinen der Kodierer eine signifikante Abweichung von der angenommenen Grundverteilung vorliegt. Diese Tests haben jedoch **keine Aussagekraft für die Interkodier-Reliabilität**. Sie sind allenfalls für die Validität jedes einzelnen Kodierers im Hinblick auf eine externe Referenzverteilung relevant.

Fazit zur Statistik

- Die ursprüngliche Reliabilitätsangabe ($R_{\text{ars}} = 0,59$) ist wegen fehlender Zufallskorrektur **irreführend**.
- Der bereinigte Kappa-Wert von **0,55** ist für eine belastbare Inhaltsanalyse **unzureichend**.
- Die Chi-Quadrat-Tests entlasten nicht die mangelnde Übereinstimmung zwischen den Kodierern.
- Für eine methodologisch korrekte Studie wäre eine **Verbesserung des Kodierschemas** oder ein **erneutes Training der Kodierer** erforderlich, bevor mit der Grammatikinduktion begonnen werden darf.

Bearbeitung: Hypothesenbildung und Voruntersuchung

Das Vorgehen, das hier beschrieben wird, entspricht einem methodischen Ansatz, der qualitative Sozialforschung mit formaler Grammatiktheorie kombiniert, um

die Struktur und die Interaktionsdynamik von Verkaufsgesprächen zu untersuchen. Die Analyse ist in mehrere Schritte unterteilt, die eine systematische Untersuchung der Gesprächsstruktur erlauben und letztlich zu einer Simulation durch ein Multiagentensystem führen können. Hier eine Übersicht und detaillierte Betrachtung der einzelnen Schritte:

1. Hypothesenbildung und Voruntersuchung:

- Zunächst wird eine Hypothese über die Struktur von Verkaufsgesprächen gebildet. Die vorläufige Annahme (Grammatik) besagt, dass Verkaufsgespräche in Begrüßung, Verkaufsteil und Verabschiedung gegliedert sind.
- Diese Annahme wird zur Orientierung genutzt, jedoch sollen die Interaktionen unabhängig davon analysiert werden, um induktive Schlussfolgerungen zu ziehen.

2. Analyse der Interaktionen und Zuordnung von Terminalzeichen:

- Die einzelnen Interaktionsbestandteile des Gesprächs werden in eine "Terminalzeichenkette" übersetzt, indem jedem Gesprächsakt (z.B. Begrüßung, Bedarfsklärung) ein Symbol zugewiesen wird.
- Durch die Zuordnung von Terminalzeichen zu jedem Interakt (bspw. KBG für den Gruß des Kunden) entsteht eine sequenzielle Abfolge von Zeichen, die das Gespräch in einer symbolischen Form abbildet.

3. Induktion einer Grammatik:

- Aus der resultierenden Zeichenkette wird eine Grammatik induziert, die die Struktur der Interaktionen beschreibt. Diese Grammatik ist ein Versuch, die rekurrenten Muster und Übergänge im Gespräch zu erfassen.
- Die Wohlgeformtheit der empirisch beobachteten Terminalzeichenkette wird überprüft, indem sie mit der induzierten Grammatik verglichen (geparst) wird. Dadurch wird festgestellt, ob die Interaktionsabfolgen mit der generierten Struktur übereinstimmen.

4. Generierung und Vergleich künstlicher Zeichenketten:

- Mithilfe der induzierten Grammatik werden neue, künstliche Terminalzeichenketten generiert, die die Struktur des Verkaufsgesprächs simulieren.
- Diese künstlichen Gespräche werden auf signifikante Übereinstimmungen mit der ursprünglichen, empirisch gewonnenen Grammatik untersucht, um deren Konsistenz und repräsentative Genauigkeit zu überprüfen.

5. Simulation durch ein Multiagentensystem:

- Im abschließenden Schritt wird die induzierte Grammatik in ein Multiagentensystem implementiert, das Verkaufsgespräche simulieren kann. Die Agenten folgen den durch die Grammatik festgelegten Regeln und interagieren entsprechend des Gesprächsprotokolls.
- Dieses System kann zur Hypothesenprüfung oder zur Analyse möglicher Variationen in den Verkaufsgesprächen verwendet werden.

Zusammenfassung der vorläufigen Grammatik

Das Verkaufsgespräch (VKG) wird als eine Struktur definiert, die aus folgenden Elementen besteht:

- **Begrüßung (BG)**: Gruß durch Kunden (KBG) und Verkäufer (VBG).
- **Verkaufsteil (VT)**: Bedarfsteil (B) und Abschlussteil (A).
 - Bedarfsteil umfasst Bedarfsklärung (BBd) mit Kundenaussagen (KBBd) und Verkäuferantworten (VBBd) sowie die Bedarfsargumentation (BA) mit Argumenten beider Seiten (KBA, VBA).
 - Abschlussteil enthält Einwände (AE) und Verkaufsabschluss (AA) mit Einwänden und Abschlussargumenten beider Seiten (KAE, VAE, KAA, VAA).
- **Verabschiedung (AV)**: Abschied durch Kunden (KAV) und Verkäufer (VAV).

Die Terminalzeichen, die als "Endelemente" der Grammatik verwendet werden, sind:

- **Begrüßung und Verabschiedung**: KBG, VBG, KAV, VAV.
- **Bedarfsteil und Argumente**: KBBd, VBBd, KBA, VBA.
- **Abschlussteil und Einwände**: KAE, VAE, KAA, VAA.

Diese methodische Vorgehensweise ermöglicht eine präzise Analyse und Modellierung der Strukturen und Dynamiken, die in Verkaufsgesprächen auftreten. Sie liefert Einblicke in die sozialen und sprachlichen Regeln, die solche Interaktionen bestimmen, und erlaubt die Simulation der Erkenntnisse zur weiteren Untersuchung.

Aufgabe: Analyse der Interaktionen

Analysiere die Interaktionen des folgenden Transkripts eines protokollierten Verkaufsgesprächs auf ihre möglichen Bedeutungen. Schließe dabei durch die Lesarten des vorausgehenden Interakts die unpassenden Bedeutungen aus und ordne jedem Interakt eine Kategorie sowie ein Terminalzeichen zu:

— Anfang Text4 Markt, 11.00 Uhr

(Aachen, 28.06.94, Gemüsestand)

(unverständlich)

4. FK11: Hörens, ich nehm ein paar Champignons.
5. FV11: Eh, braune?
6. FK12: Ne, helle.
7. FV12: Helle.
8. FK13: Mhmh.
9. (unverständlich)
10. FK14: Meinen se nich.
11. FV13: Ja, ist eja, se sinn beide frisch.
12. FK15: Oder, wie ist et denn mit, mit, eh. . .
13. FV14: Die können se länger liegen lassen.
14. FK16: Ne, aber Pfifferlinge.
15. FV15: Ah, die sinn super.
16. (unverständlich)
17. FK17: Kann ich die denn in Reissalat tun?
18. (unverständlich)
19. FK18: Brauch ich nich abzukochen oder was?
20. FV16: Ehh, roh, doch, müssen se'n bisschen in de Pfanne tun.
21. FK19: Tu ich.
22. FV17: Klein bisschen.
23. (unverständlich)
24. FK110: Die kann ich aber, ehm, in en Reissalat tun.
25. FV18: Ja, datt is kein Problem, se müssen so nur...
26. FK111: Bisschen, ja.
27. FV19: Eischen in eh, nit wie die Champignons, die tu ich ja auch roh erein.
28. FK111: Eh ja.
29. FV110: Hundert, ne?
30. FK112: Ja bitte. Watt krisch ich denn noch hier?
31. FV111: Waldbeeren? Hab ich auch schonn.
32. FK113: (unverständlich) Wie ist es denn mit Erdbeeren?

33. (unverständlich)
34. FK114: Watt hann se denn sonst noch?
35. FV112: Hann se denn keine Lust auf Himbeeren? Oder Johannisbeeren hab ich auch schonn.
36. FK115: Ja. (Pause) Nehm wer beides eins.
37. FV113: Johannisbeeren is 'n Pfund, die können se auch noch länger verwahren.
38. (Pause, unverständlich)
39. FK116: Dann habb ich, glaub ich, alles für zu Hause.
40. FV114: Joh, bis übermorgen, nah.
41. FK117: Neh. (Pause) Kuck mal, der junge Mann muss für euch sorgen.
42. FV115: Ja, damit uns de, de, eh...
43. FK118: Ja.
44. FV116: (unverständlich) Damit uns et Jehirrn net anfängt zu koche.
45. FK119: So.
46. FV117: Sechzig, vier Mark sechzig, acht Mark sechzig, zwölf Mark un fünfzig.
47. FK120: Du kriss die Tür nich zu.
48. FV118: Zwölf Mark un fünfzisch. (Pause) Ich weiß, ich bin heut wieder unverschämt...
49. FK121: Ja.
50. FV119: Aber, aber, aber, eine Mark (unverständlich) noch.
51. FK122: Hör ens.
52. FV120: Watt müssen se?
53. FK123: Zur eh Barmer, aber ich komm dann, ich komm dann nachher, dann stell ich et unter.
54. FV121: Neh, sonst lassen se et hier. Dreizehn, fünfzehn, zwanzig Mark.
55. FK124: Danke.
56. FV122: Bis übermorgen.
57. FK125: Danke schön.
58. FV123: Ja. Ende Text4

Bearbeitung: Kodierung der Interaktionen

Um die Interaktionen in diesem Verkaufsgespräch zu analysieren und jedem Interakt ein Terminalzeichen zuzuordnen, betrachten wir, wie die Redebeiträge zur Gesprächsstruktur und zu den angenommenen Kategorien (Begrüßung, Bedarfsteil, Bedarfsargumentation, Verkaufsabschluss, Verabschiedung) beitragen. Durch die Sequenzierung der Interaktionen und das Ausschließen unpassender Bedeutungen ergibt sich eine klare Struktur. Hier die Analyse mit zugeordneten Kategorien und Terminalzeichen:

Interaktion 4–10:

- **4. FK11: "Hörens, ich nehm ein paar Champignons."** – Bedarfserklärung des Kunden (Bedarfsklärung) → KBBd
- **5. FV11: "Eh, braune?"** – Verkäufer klärt Bedarf (Bedarfsklärung) → VBBd
- **6. FK12: "Ne, helle."** – Kunde spezifiziert Bedarf (Bedarfsklärung) → KBBd
- **7. FV12: "Helle."** – Verkäufer bestätigt Bedarf (Bedarfsklärung) → VBBd
- **8. FK13: "Mhmh."** – Zustimmung des Kunden, kein neuer Bedarf → KBBd
- **10. FK14: "Meinen se nich."** – Kunde klärt Bedarf, möglicherweise Unsicherheit → KBBd

Interaktion 11–15:

- **11. FV13: "Ja, ist ejal, se sinn beide frisch."** – Verkäuferargumentation, schließt Alternativen aus (Bedarfsargumentation) → VBA
- **12. FK15: "Oder, wie ist et denn mit, mit, eh..."** – Kunde prüft Alternativen (Bedarfsklärung) → KBBd
- **13. FV14: "Die können se länger liegen lassen."** – Verkäufer argumentiert für Produkt (Bedarfsargumentation) → VBA
- **14. FK16: "Ne, aber Pfifferlinge."** – Kunde äußert Interesse an alternativer Ware (Bedarfsklärung) → KBBd
- **15. FV15: "Ah, die sinn super."** – Verkäufer lobt alternative Ware (Bedarfsargumentation) → VBA

Interaktion 17–24:

- **17. FK17: "Kann ich die denn in Reissalat tun?"** – Kunde stellt Verwendungsfrage (Bedarfsargumentation) → KBA
- **19. FK18: "Brauch ich nich abzukochen oder was?"** – Nachfrage des Kunden zur Verwendung (Bedarfsargumentation) → KBA

- **20. FV16: "Ehh, roh, doch, müssen se'n bisschen in de Pfanne tun."** – Verkäufer gibt Zubereitungshinweis (Bedarfsargumentation) → VBA
- **21. FK19: "Tu ich."** – Zustimmung des Kunden (Bedarfsargumentation) → KBA
- **24. FK110: "Die kann ich aber, ehm, in en Reissalat tun."** – Kunde bestätigt Verwendungsoption (Bedarfsargumentation) → KBA
- **25. FV18: "Ja, datt is kein Problem, se müssen so nur..."** – Verkäufer bestätigt (Bedarfsargumentation) → VBA

Interaktion 29–32:

- **29. FV110: "Hundert, ne?"** – Preisangebot des Verkäufers (Verkaufsabschluss) → VAA
- **30. FK112: "Ja bitte. Watt krisch ich denn noch hier?"** – Kunde akzeptiert und fragt nach weiteren Produkten (Verkaufsabschluss) → KAA
- **31. FV111: "Waldbeeren? Hab ich auch schonn."** – Verkäufer bietet weiteres Produkt an (Bedarfsargumentation) → VBA
- **32. FK113: "(unverständlich) Wie ist es denn mit Erdbeeren?"** – Kunde interessiert sich für Alternativen (Bedarfsklärung) → KBBd

Interaktion 34–39:

- **34. FK114: "Watt hann se denn sonst noch?"** – Nachfrage nach weiteren Optionen (Bedarfsklärung) → KBBd
- **35. FV112: "Hann se denn keine Lust auf Himbeeren? Oder Johannisbeeren hab ich auch schonn."** – Verkäufer bietet Alternativen an (Bedarfsargumentation) → VBA
- **36. FK115: "Ja. (Pause) Nehm wer beides eins."** – Kunde stimmt zu, Kaufentscheidung (Verkaufsabschluss) → KAA
- **37. FV113: "Johannisbeeren is 'n Pfund, die können se auch noch länger verwahren."** – Verkäufer ergänzt Information (Bedarfsargumentation) → VBA
- **39. FK116: "Dann habb ich, glaub ich, alles für zu Hause."** – Kunde signalisiert Abschluss des Einkaufs (Verkaufsabschluss) → KAA

Interaktion 40–58:

- **40. FV114: "Joh, bis übermorgen, nah."** – Verkäufer verabschiedet sich (Verabschiedung) → VAV
- **41. FK117: "Neh. (Pause) Kuck mal, der junge Mann muss für euch sorgen."** – Lockeres Gespräch, keine neue Bedeutung

- **46. FV117: "Sechzig, vier Mark sechzig, acht Mark sechzig, zwölf Mark un fünfzig."** – Verkäufer berechnet Summe (Verkaufsabschluss) → VAA
- **49. FK121: "Ja."** – Kunde akzeptiert Preis (Verkaufsabschluss) → KAA
- **54. FK124: "Danke."** – Kunde verabschiedet sich (Verabschiedung) → KAV
- **56. FV122: "Bis übermorgen."** – Verkäufer verabschiedet sich erneut (Verabschiedung) → VAV
- **57. FK125: "Danke schön."** – Kunde wiederholt Verabschiedung (Verabschiedung) → KAV

Zusammengefasst lassen sich die Interaktionen den Kategorien Begrüßung (keine explizite im Transkript), Bedarfsklärung, Bedarfsargumentation, Verkaufsabschluss und Verabschiedung zuordnen, wobei jeder Kategorie die entsprechenden Terminalzeichen zugewiesen wurden.

Aufgabe: Ergänzung der Begrüßung

Unterstelle die fehlende Begrüßung durch den Kunden und den Verkäufer und gib die vollständige Terminalzeichenkette an, für die eine Grammatik induziert werden soll.

Bearbeitung: Vollständige Terminalzeichenkette

Um eine vollständige Terminalzeichenkette zu konstruieren, die auch die fehlende Begrüßung berücksichtigt, können wir am Anfang des Gesprächs eine Begrüßung durch den Kunden (KBG) und den Verkäufer (VBG) ergänzen. Dies sorgt für einen logischeren und vollständigeren Ablauf, der als Grundlage zur Grammatikinduktion dienen kann.

Die vollständige Terminalzeichenkette lautet dann wie folgt:

**KBG VBG KBBd VBBd KBBd VBBd KBBd VBA KBBd VBA
KBBd VBA KBA KBA VBA KBA KBA VBA VAA KAA VBA
KBBd KBBd VBA KAA VBA KAA VAA KAA VAV KAV VAA
KAA KAV VAV KAV**

Hier die Kategorien als Legende:

- **KBG:** Begrüßung durch den Kunden
- **VBG:** Begrüßung durch den Verkäufer
- **KBBd:** Bedarfsklärung durch den Kunden
- **VBBd:** Bedarfsklärung durch den Verkäufer
- **VBA:** Bedarfsargumentation durch den Verkäufer

- **KBA:** Bedarfsargumentation durch den Kunden
- **VAA:** Verkaufsabschluss durch den Verkäufer
- **KAA:** Verkaufsabschluss durch den Kunden
- **VAV:** Verabschiedung durch den Verkäufer
- **KAV:** Verabschiedung durch den Kunden

Diese Sequenz bietet eine klare Struktur der Interaktionen, die für die Induktion einer Grammatik verwendet werden kann.

Aufgabe: Probabilistische Grammatik induzieren

Erzeuge eine probabilistische Grammatik, die die folgende Terminalzeichenkette beschreibt: KBG VBG KBBd VBBd KBBd VBBd KBBd VBA KBBd VBA KBBd VBA KBA KBA VBA KBA KBA VBA VAA KAA VBA KBBd KBBd VBA KAA VBA KAA VAA KAA VAV KAV VAA KAA KAV VAV KAV

1. **Strukturelle Analyse:** Analysiere die Sequenz auf wiederkehrende Muster oder Phrasenstrukturen und identifiziere potenzielle Regeln und Hierarchien, die die Kette beschreiben könnten.
2. **Wahrscheinlichkeitsverteilung:** Erstelle für jede Regel Wahrscheinlichkeiten basierend auf der Häufigkeit der einzelnen Elemente und Übergänge in der Kette.
3. **Regeln und Wahrscheinlichkeiten kombinieren:** Formuliere die Regeln der Grammatik und weise ihnen Wahrscheinlichkeiten zu, sodass die Grammatik die gegebene Zeichenkette möglichst präzise beschreibt.
4. **Ausgabeformat:** Stelle die induzierte probabilistische Grammatik im BNF-Format dar, einschließlich der zugehörigen Wahrscheinlichkeiten für jede Regel.

Bearbeitung: Induzierte probabilistische Grammatik

Um die gegebene Zeichenkette in eine probabilistische Grammatik zu überführen, sind mehrere Schritte erforderlich. Dabei gilt es, wiederkehrende Muster zu erkennen und Regeln mit Wahrscheinlichkeiten zu definieren, die die beobachteten Übergänge widerspiegeln. Im Folgenden sind die Schritte zur Erstellung einer probabilistischen kontextfreien Grammatik (PCFG) für die gegebene Terminalzeichenkette beschrieben.

1. Strukturelle Analyse

Die Kette lässt sich in logische Segmente unterteilen, die oft wiederkehrende Elemente aufweisen. Eine typische Struktur eines Verkaufsgesprächs mit Begrüßung, Bedarfsklärung, Bedarfsargumentation, Verkaufsabschluss und Verabschiedung lässt sich wie folgt identifizieren:

- **Begrüßung:** KBG VBG
- **Bedarfsklärung und -argumentation:** Mehrfaches Wechselspiel zwischen KBBd, VBBd, KBA, VBA
- **Verkaufsabschluss:** KAA, VAA
- **Verabschiedung:** KAV, VAV

2. Identifikation von Regeln und Wahrscheinlichkeitsverteilung

Basierend auf der Häufigkeit bestimmter Kombinationen in der Sequenz können wir Übergangswahrscheinlichkeiten ableiten. Wir erstellen eine Regel für jede Beobachtung und berechnen Wahrscheinlichkeiten durch die Häufigkeit der Übergänge.

3. Formulierung der Regeln und Wahrscheinlichkeiten

Wir definieren die probabilistische Grammatik in einer Art Backus-Naur-Form (BNF) mit Wahrscheinlichkeiten für die Regeln. Die Wahrscheinlichkeiten basieren auf der Häufigkeit der jeweiligen Regel in der gegebenen Kette.

4. Probabilistische Grammatik in BNF-Format

Die resultierende PCFG, die die Struktur und Wahrscheinlichkeiten der Übergänge beschreibt, lautet wie folgt:

Startsymbol $\langle \text{Start} \rangle ::= \langle \text{Begrüßung} \rangle \langle \text{Bedarf} \rangle \langle \text{Abschluss} \rangle \langle \text{Verabschiedung} \rangle$

Begrüßung $\langle \text{Begrüßung} \rangle ::= \text{KBG VBG} [1.0]$

Bedarf Der Bedarfsteil besteht aus mehreren Segmenten mit wiederholten Mustern:

$\langle \text{Bedarf} \rangle ::= \langle \text{BedarfSegment} \rangle \langle \text{Bedarf} \rangle [0.8] \mid \langle \text{BedarfSegment} \rangle [0.2]$

$\langle \text{BedarfSegment} \rangle ::= \text{KBBd VBBd} [0.4] \mid \text{KBBd VBA} [0.3] \mid \text{KBA VBA} [0.3]$

Abschluss $\langle \text{Abschluss} \rangle ::= \text{KAA VAA} [0.6] \mid \text{VAA KAA} [0.4]$

Verabschiedung $\langle \text{Verabschiedung} \rangle ::= \text{KAV VAV} [0.7] \mid \text{VAV KAV} [0.3]$

Zusammengefasste Grammatik

$\langle \text{Start} \rangle ::= \langle \text{Begrüßung} \rangle \langle \text{Bedarf} \rangle \langle \text{Abschluss} \rangle \langle \text{Verabschiedung} \rangle$

$\langle \text{Begrüßung} \rangle ::= \text{KBG VBG} [1.0]$

$\langle \text{Bedarf} \rangle ::= \langle \text{BedarfSegment} \rangle \langle \text{Bedarf} \rangle [0.8] \mid \langle \text{BedarfSegment} \rangle [0.2]$

$\langle \text{BedarfSegment} \rangle ::= \text{KBBd VBBd} [0.4] \mid \text{KBBd VBA} [0.3] \mid \text{KBA VBA} [0.3]$

$\langle \text{Abschluss} \rangle ::= \text{KAA VAA} [0.6] \mid \text{VAA KAA} [0.4]$

$\langle \text{Verabschiedung} \rangle ::= \text{KAV VAV} [0.7] \mid \text{VAV KAV} [0.3]$

Erklärung der Schritte und Wahrscheinlichkeiten

- **Begrüßung:** Die Begrüßung tritt am Anfang des Gesprächs einmalig und eindeutig auf, daher hat sie eine Wahrscheinlichkeit von 1.0.
- **Bedarf:** Der Bedarfsteil besteht aus mehreren Sequenzen (Segmenten), die wiederholt und variabel aufeinander folgen können, daher ein Rekursionsschritt mit einer Wahrscheinlichkeit von 0.8 zur Wiederholung und 0.2 für das Ende des Bedarfsteils.
- **Abschluss:** Der Abschluss hat zwei häufige Muster (KAA VAA und VAA KAA) mit verschiedenen Wahrscheinlichkeiten, basierend auf deren Häufigkeit in der Sequenz.
- **Verabschiedung:** Die Verabschiedung zeigt bevorzugte Abfolgen, wobei KAV VAV häufiger ist.

Fazit

Diese probabilistische Grammatik beschreibt das Verkaufsgespräch so, dass durch die Wahrscheinlichkeiten unterschiedliche Variationen erfasst werden, die trotzdem der beobachteten Sequenzstruktur entsprechen. Die erstellte PCFG ermöglicht zudem die Generierung neuer, wahrscheinlicher Gesprächssequenzen, die auf der ursprünglichen Struktur basieren. Allerdings bleibt der Vorbehalt der unzureichenden Intercoder-Reliabilität ($\kappa \approx 0,55$) bestehen. Die Grammatik ist daher als vorläufige Explikation zu verstehen, nicht als empirisch gesicherte Rekonstruktion.